FINAL REPORT

19960415

A GENOMICS-BASED PROGRAM FOR BRASSICA CROP DEVELOPMENT

November 2002
Funded by the Agri-Food Innovation Fund

Prepared by: Plant Biotechnology Institute
& Ag and Agri-Food Canada

AGRI-FOOD INNOVATION FUND

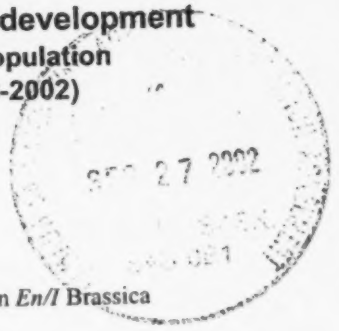
Canada

Saskatchewan

A genomics-based program of Brassica crop development

Generation of a *Brassica napus En/I* tagged Population

AFIF program Final Report (PBI Participants; 1998-2002)



Specific objective(s)

Generation of En/I tagged Brassica founder lines

- Analysis of active and independent transposition events in *En/I* Brassica populations and subsequent generations
- Determination of transposition frequency

Milestones

Year 1

- T-DNA vectors containing optimised *En*- and *I*- elements designed and constructed.

Year 2

- Successful transformation of the selected recipient line of *B. rapa* with *En*- and *I*- elements

Year 3

- At least 15 transformants containing *En* constructs and at least 30 transformants containing *I*- elements constructs
- Transposition of *I* elements demonstrated in *B. rapa*.

Year 4

- At least 15 transformants containing *En* constructs and at least 30 transformants containing *I- elements* constructs
- Good molecular and genetic characterization of the transposition process
- Efficient maintenance and multiplication of the mutagenetic base population

The project will make possible the generation of a population carrying thousands of distinct transposition insertion events, starting from a relatively small number of transformants as founder lines expressing the active form of transposase in combination with mobilizable *I*- elements.

Progress (Jan 1998-Aug 2002)

Large scale mutagenesis projects are tools for providing useful agronomic phenotypes, and are powerful tools for the elucidation of gene function when used in combination with reverse genetics strategies.

Brassica napus is a plant of agronomic and economical importance for the Canadian prairies. The development of a mutagenized *Brassica napus* population will allow the selection of mutations that cause subtle changes in the phenotype and allow mutations to be selected in

pairs of duplicate genes. The population will yield totally to novel genetic variation for use in future crop improvement.

Two approaches are generally available for the conduct of large scale mutagenesis in plants: T-DNA tagging and transposon tagging. Both methods have proven to be very efficient, however transposon tagging is the method of choice when the studied system has a low frequency of transformation such as *Brassica napus*. Indeed, a relatively small set of selected transposon-tagged founder lines can be propagated to generate a large population containing many different transposition insertion events.

Selection of plant material and T-DNA constructs

Brassica napus was chosen in this study instead of *Brassica rapa* (which was proposed in the original project) due to its growing economic importance to Canadian agriculture. The elite line DH-12075 (G. Rakow, AAFC) was selected because of the many desirable agronomic traits evident in this line. The T-DNA construct *cwEn/I* (Aarts *et al* (1995); kindly provided by A. Pereira, Wageningen, Netherlands; Fig 1) contained *En/I* transposable elements from *Zea mays* (Schlappi *et al,* 1993) and was chosen and modified in preparation for Brassica transformation.

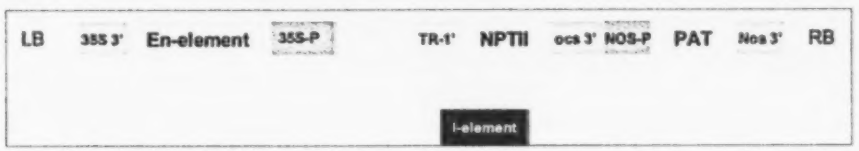


Fig 1. Schematic Diagram of the *cwEn/I* Construct

For better control and enhanced frequency of the transposition event, the expression construct used to express the transposase (enzyme needed for transposition) was transformed independently from the construct containing the mobile element (transposon). Accordingly, *En* has its terminal inverted repeats removed and was used only as a source of transposase. The non autonomous *I element* was the only construct capable of transposition and was inserted downstream of the ATG codon of the *NPTII* gene. The *NPTII* gene was used as an excision marker, whereas the *PAT* gene (which confers the resistance to the herbicide phosphinotricin [PPT]) was used as a transformation marker.

Agrobacterium*-mediated transformation of *Brassica napus

The T-DNA construct was introduced into *Agrobacterium tumefaciens* strain GV3101 containing the helper plasmid construct pMP90RK. A suspension of recombinant *Agrobacterium* was used to transform hypocotyl segments by co-cultivation according to Katavic *et al.* (1996). The transformation frequency was assessed at approximately 5% of viable explants. The sequence of transformation steps and their chronology are diagramatically summarized in Fig.2.

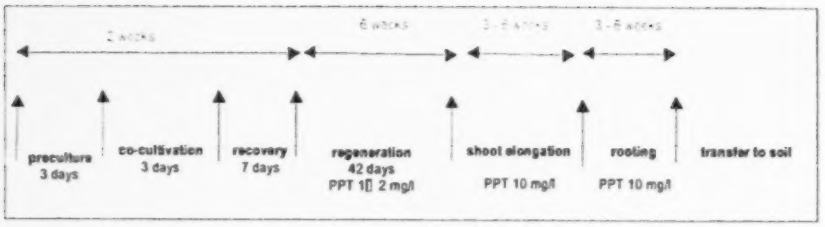


Fig.2: Diagram of the Brassica transformation protocol. Primary transformants (T0) placed to soil, grown under standard greenhouse conditions and seeds harvested.

Approximately 200 *En/I* tagged *B. napus* founder lines were identified and selected based on 3 criteria: i) copy number of T-DNA inserted in the genome, ii) expression of the transposase gene and iii) number of seeds recovered. Each of these selected plants were self-pollinated and progated to generate the T3 population.

Analyses of transpositions event among *En/I* primary transformants, T1 and T2

Two non lethal methods were used assess the molecular basis and frequency of transposition:

1. *NPTII* enzymatic activity: this method presented two drawbacks, including an unacceptably high threshold, coupled with the finding that only 1/3 of transpositions could only be detected (the ones which restored the *NPTII* reading frame)
2. A PCR based method, using primers which amplified the donor site following departure of the *I element*. This method allowed the detection of all transpositions regardless the *NPTII* reading frame and, when combined with sequencing, resulted in well-characterized 'footprints' left by the *I element* after excision at the donor site. This method allowed us to detect multiple transposition events within the same plant

The results of sequence-characterization of plants selected by PCR for departure of the *I-element* are summarized in Fig 3.

```
366-1   GTGTGCTGGA ATTCGCCCTT AAAAAATTCA TATCTTCCAC ACGCGATGGA
366-2   GTGTGCTGGA ATTCGCCCTT AAAAAATTCA TATCTTCCAC ACGCGATGGA
366-3   GTGTGCTGGA ATTCGCCCTT AAAAAATTCA TATCTTCCAC ACGCGATGGA
366-4   GTGTGCTGGA ATTCGCCCTT AAAAAATTCA TATCTTCCAC ACGCGATGGA

366-1   TCCGGCCTGT CCAAAGACCC TGATCAGGTA GATCCGGCCG GATCCGGCCA
366-2   TCCGGCCTGT CCAAAGACCC TGA---GGTA GATCCGGCCG GATCCGGCCA
366-3   TCCGGCCTGT CCAAAGAC-- TGA---GGTA GATCCGGCCG GATCCGGCCA
366-4   TCCGGCCTGT CCAAAGACCC TGA------A GATCCGGCCG GATCCGGCCA

366-1   AGCTAGCTTG GATTGAACAA GATGGATTGC ACGCAGGTTC TCCGGCCGCT
366-2   AGCTAGCTTG GATTGAACAA GATGGATTGC ACGCAGGTTC TCCGGCCGCT
366-3   AGCTAGCTTG GATTGAACAA GATGGATTGC ACGCAGGTTC TCCGGCCGCT
366-4   AGCTAGCTTG GATTGAACAA GATGGATTGC ACGCAGGTTC TCCGGCCGCT

Tr'1 promoter
In frame with NPTII
NPTII gene
```

Fig.3: Summary of sequence characterization of *I-element*"footprints"in select lines of Brassica founder lines

The overall frequency of transposition among select founder lines was estimated to be approximately 41%, as summarized below in Table 1.

	Transposed lines	Analysed lines	Frequency of transposition
1 copy of T-DNA	14	73	19%
2 copies of T-DNA	23	45	51%
more than 2 copies of T-DNA	39	69	57%
Total	76	187	Mean 41%

Table 1: Estimated Frequency of transposition among founder lines of *B. napus*

The overal transposition frequency was observed to increase in correlation with the number of T-DNAs inserted to the. However, further analyses showed that all transposition events were restricted to the primary transformants, and an early stage in, and were transmitted to the T1 and T2 populations.

Improvement of transposition frequency by manipulation of *tnpA* splice variant expression

In *Maize*, the *En* transposon encodes four different mRNA transcripts via alternative splicing: *tnpA*, *tnpB*, *tnpC* and *tnpD*. Both *tnpA* and *tnpD* have been shown to be essential for the transposition event when expressed at an optimal ratio 100:1. Our own molecular studies indicated that the presence of the *tnpA* transcript variant from the native maize *En* transposon is very low in transgenic Brassica, and that the *tnpA*/*tnpD* ratio is inverted (1:100) in comparison withother transgenic plant species reported in the literature-probably due to a fundamental difference in RNA splicing capacity.

In order to improve the usefulness of the population and increase the frequency of transposition, we undertook to increase the level of *tnpA* transcripts in transgenic DH-12075. Two strategies were used:

- Secondary (re)transformation of selected *En/I* lines (T3 generation) with a *tnpA* overexpression construct
- Cross-pollination of the existing *En/I*-containing population with DH-12075 recombinant lines carrying the *tnpA* overexpression construct.

Results for the expression of *tnpA* cDNA for double-recombinant *B. napus lines* recovered by cross-pollination with the existing population are summarized in Fig 4.

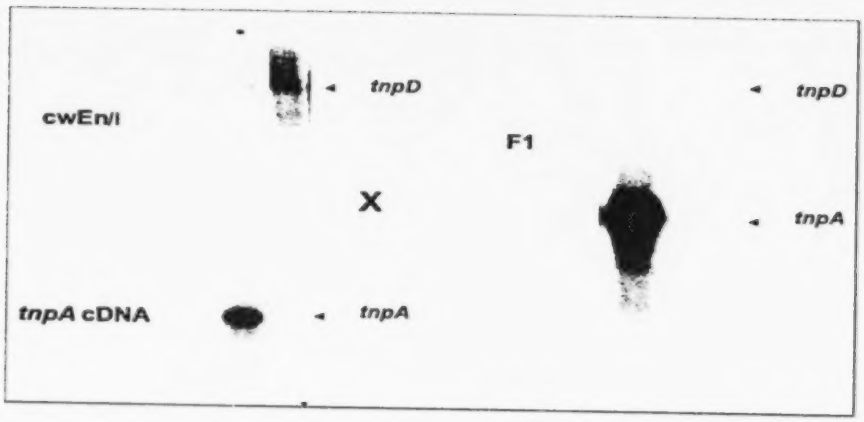


Fig 4: Analysis of *tnpA* and *tnpD* transcripts in F1 population after cross-pollination between *tnpA* cDNA and *En/I* populations.

Among the F1 population analyzed, *tnpA* transcript levels were found to be increased and the *tnpA/tnpD* ratio for efficient transposition was restored. In this study, 25 F1 plants derived from different crosses were studied for transposition events, where preliminary data showed that one of these carried an independent transposition event.

In a second series of experiments, selected *En/I* lines (the T3 generation) were re-transformed witha *tnpA* cDNA construct, and the primary retransformants (T0) were propagated in the greenhouse to yield a T1population. Plants containing both the *En/I* and *tnpA* cDNA constructs were analyzed for transposition events using the aforementioned PCR-based method amplifying the donor site after excision of *I element*.

Founder line	T1 line after retransformation	number of plants analyzed	number of plants PCR positive	Frequency of transposition %
11	rt-11-2b-()	26	10	38.5
	rt-11-2d-()	46	14	30.4
	rt-11-2f-()	62	23	37.1
	rt-11-4a-()	54	13	24.1
	rt-11-8-()	87	73	83.9
	rt-11-10-()	91	54	59.3
13	rt-13-3c-()	76	0	0.0
	rt-13-8-()	64	7	10.9
207	rt-207-4-()	48	28	58.3

Table 2: Frequency of *I-element* transposition among T1 retransformed lines carrying a *tnpA* cDNA expression construct

The frequency of transposition of different retransformed lines are summarized in Table 2. Briefly, the transposition frequency was found to vary from 0 to 83.9%, suggesting that transposition events occur at different stages of the *Brassica napus* life cycle. An early transposition would be expected to be transmitted to substantially all the progeny (high frequency of transposition), whereas late-stage transposition would be transmitted to a smaller percentage

of the progeny (low frequency of transposition). Among select lines, independent transpositions were also characterized using Southern blotting experiment using the *I element* as a probe.

Summary and Conclusions

We generated over 200 *En/I* tagged *Brassica napus* founder lines and determined their frequency of transposition. Among these lines, transposition events were found to occur only in the primary transformants and not in the subsequent generations, probably due to the low expression of *tnpA* product necessary for the transposition event. To improve the usefulness of our founder lines, we undertook to increase the level of *tnpA* product in the plants using two strategies:

- Retransformation of some *En/I* selected lines (T3 generation) with *tnpA* cDNA
- Generation of *tnpA* cDNA *Brassica napus*- Cross-pollination with the existing population.

In both approaches, the level of *tnpA* transcripts was increased and the *tnpA/tnpD* ratio for efficient transposition was restored among the transgenic and hybrid lines analyzed. Analysis of a set of plants showed the presence of active transposition, including independent transposition events among the progeny. These new populations constitute a valuable resource as founder lines for generating a large *En/I* tagged Brassica population.

Variation from the planned milestones

Year 1

- Use of a T-DNA containing *En*- and *I*- elements *in cis*. This T-DNA has been proven to be very efficient in generating a transposon tagged Arabidopsis population. The T-DNA was modified for better transformation of *Brassica napus*.
- By the fact to have both elements *in cis* on the same T-DNA, one needs only to generate one population of *En/I* tagged *Brassica napus*, instead of generating two populations containing each one of the transposable elements and cross-pollinate the primary transformants to get finally the desired *En/I* population.
- Successful transformation of DH12075 with T-DNA containing *En/I in cis*.

Year 2

- Generation and selection of 200 founder lines for further studies
- Propagation of the founder lines in the greenhouse: T1 and T2 populations
- Analyses of transposition events in subsequent generations

Year 3

- Propagation of T2 plants to T3 plants in the greenhouse
- Proof of transposition events only in primary transformants.
- Molecular analysis of transposition event. Need to have *tnpA* and *tnpD* products for transposition. Level of *tnpA* products too low in *Brassica napus*.
- Increase level of *tnpA* for efficient transposition

- Transformation of *Brassica napus* DH12075 with *tnpA* cDNA. Generation of 29 *tnpA* cDNA containing DH12075.
- Cross-pollination with selected *En/I* lines. 108 crosses performed
- Retransformation of 4 selected *En/I* lines with *tnpA* cDNA. Obtention of 11 retransformants.

Year 4

- Restoration of *tnpA/tnpD* ratio for efficient transposition in retransformed and cross-pollinated plants.
- Study for transposition in progeny of retransformed and cross-pollinated plants

AFIF funded Staff

Elizabeth Nanak, PhD (July 22/98-Feb 28/02)
Nataliya Dormann, MSc (Jan 18/98-Jan 18/02)
Alan Weninger, BSc (Jan /99-Feb 9/01)
Debra Brogden, BSc (April 1/00-Sept 30/00)
Cyril Periappuram, PhD (Feb/01-Oct 5/01)
Susan Campbell, BSc (Oct 1/00-Aug 31/01)

Outputs (Deliverables)

Plant material

En/I tagged Brassica napus founder lines are stored at TPC/PBI/NRC and are available for further investigation, upon request to:

bcrosby@gene.pbi.nrc.ca
enanak@gene.pbi.nrc.ca

Presentation and Publications

1. Nanak, E., Dormann, N., Campbell, S., Weninger, A., Periappuram, C., Brogden, D. and Crosby, W. Generation and molecular analyses of *Brassica napus En/I* tagged population. *Article in preparation*
2. Nanak, E., Dormann, N., Campbell, S., Weninger, A., Periappuram, C., Brogden, D. and Crosby, W. Generation of *Brassica napus En/I* tagged population. Canadian Plant Tissue Culture and Genetic Engineering Workshop, Saskatoon, SK, Canada, 12-15 May 2001. *Conference presentation*
3. Nanak, E., Dormann, N., Campbell, S., Weninger, A., Periappuram, C., Brogden, D. and Crosby, W. A genomics-based program for Brassica crop development. Canada-Saskatchewan Agri-food innovation fund conference, Regina, SK, Canada, 7-8 Nov 2000. *Poster Presentation*

Literature Cited

1. Aarts, M.C.M., Keijzer, C.J., Stiekema, W.J., Pereira, A. (1995), Mol. Gen. Genet. 247, 555-564
2. Katavic, V., Campbell, S., Friesen, L., Palmer, D., Keller, W., Taylor, D. (1996), 4th Plant Tissue Culture & Genetic Engineering Conference, 103
3. Schlappi, M.R., Smith, D., Fedoroff, N. (1993), Genetics 133:1009-1021

103-03569/

Final Report for AFIF grant (AAFC-SRC contribution)

Date: April 9th, 2002

Reporting Period: Sept 31st, 2001 to March 31st, 2002.

Principal Investigator: Dr Derek Lydiate

Project Title: **A Genomics-Based Program for *Brassica* Crop Development**

Abstract:

It is becoming increasingly possible to describe aspects of basic plant biology in terms of the integrated actions of defined genes. Genome sequencing initiatives are accelerating this process and will eventually lead to vast numbers of genes with defined physiological, biological and/or developmental roles being precisely localized on the genomes of model plants. Crop scientists and plant breeders working with *Brassica* species stand to gain a tremendous amount from the new possibilities arising from this explosion of knowledge since the model dicotyledonous species, *Arabidopsis thaliana* is a close relative of *Brassica* species.

There are two main hurdles to the exploitation of knowledge and resources developed in model systems, (1) identifying genes in the model species equivalent to genes controlling useful genes in the crop plants and (2) developing or identifying variation in crop plants in the genes considered to be of importance in the model system. The main objectives of the research programme were to address these two issues, chart the exact pattern of collinearity between the genomes of cultivated *Brassica* species and that of the model dicot *A. thaliana* (AAFC-SRC contribution) and use a heterologous transposon system to develop a highly mutagenised population of *B. napus* (NRC-PBI contribution).

Executive summary:

Over the course of the project the main objectives assigned to AAFC-SRC have been achieved, 1) A reference population for genetic mapping in *B. napus* has been developed and is already being requested for use by international collaborators. 2) A comprehensive comparative map describing the regions of homology between *B. napus* and *A. thaliana* has been developed. 3) A *B. napus* large insert library for physical mapping in Brassica species has been developed and is being actively characterized. Some limited work has been carried out for the mapping of the insertion sites of I elements positioned in the mutagenised *B. napus* population, but the problems encountered with the development of the population has delayed this process. However, AAFC-SRC will continue to collaborate with NRC-PBI beyond the term of the project in order to assist in the mapping of I elements.

Technical Report:

Comparative Mapping between *Brassica napus* and *Arabidopsis thaliana.*

A global comparison between the genomes of *B. napus* and *A. thaliana* has been completed, which shows the relationship between over 1200 data points on the *B. napus* genome and over 500 data points on the *A. thaliana* genome. Presently a web site is being developed to display the data and allow easy access to interested Brassica researchers. This work has been used to establish the regions of homology between *B. napus* and *A. thaliana* for areas of the *B. napus* genome containing both a blackleg resistance gene on *B. napus* linkage group N7 and a turnip mosaic virus resistance gene on *B. napus* linkage group N6. Further fine mapping of these targeted regions has identified intra-chromosomal duplications on both linkage groups N6 and N7, such rearrangements may be a common feature of Brassica genomes and could lead to mapping errors when attempting to identify the homologous physical region in *B. napus*. Collinearity in these regions has also allowed the identification of further markers for fine and physical mapping of these two regions and in both cases there are a number of potential candidate genes which are being further studied to investigate whether they play a role in the disease resistance phenotype.

Physical Mapping of Targeted regions of the *Brassica napus* genome.

Physical mapping of a number of regions of the *B. napus* genome is underway. To test the coverage of the BAC library and to assess difficulty of identifying and aligning contigs within an amphidiploid genome, physical contigs homologous to a 2 Mb region of *A. thaliana* chromosome five are being developed. Markers separated at 50 Kb intervals along the region in *A. thaliana* have been identified; either Brassica cDNA clones or primers were designed to predicted genes identified from the *A. thaliana* genomic sequence. Of the 40 probes, 15 have been used thus far to screen the *B. napus* BAC library, 113 clones have been identified and DNA has been isolated for these clones and DNA fingerprints for each clone have been developed. This work is in progress and will be completed as part of the AAFC Genomics program. Physical maps corresponding the region of linkage groups N6 and N7, which were described above in the comparative mapping section, are also being developed. In this case the focus is only on the N6 and N7 region respectively and not on developing contigs for their duplicated copies found within the *B. napus* genome. In each case the resistance gene has been fine mapped using a set of over 1000 recombinant *B. napus* lines to a region of the genetic map corresponding to 1cM. Markers mapped to the region have been used to screen the *B. napus* BAC library, presently positive confirmation of the clones originating from N6 and N7 is underway. This involves isolating DNA from the individual BAC clones, digesting the BAC clones with the restriction enzyme which identified the original RFLP polymorphism and probing a southern filter of the digested BACs with the probe used to isolate them. Each of the BAC clones has also been fingerprinted by restriction digestion with a single enzyme, resolution on an agarose gel and fluorescent imaging upon exposure to the dye SyberGreen. This work is in

progress, the N6 study will be completed as part of the AAFC Genomics program and the N7 study will be completed as part of a collaboration with Dr Allen Good at the U of Alberta funded by an AARI grant.

Mapping of I-element insertion sites

Although the reference *B. napus* genetic mapping population is available to position the I element insertion sites, no flanking sequence has been available to use as a probe in the mapping studies. However, the work to clone and identify this sequence is on-going at NRC-PBI and AAFC-SRC will, when that sequence becomes available, map the position of the insertion sites within the *B. napus* genome.

Project Developed Materials:

A doubled haploid reference mapping population (SG population) of 300 perfectly homozygous *B. napus* lines derived from a highly polymorphic cross between DH12075 (an established canola line) and a newly resynthesised *B. napus* line.

A genetic linkage map for the SG population consisting of 321 RFLP derived loci and 300 microsatellite marker derived loci.

A comprehensive comparative map of *Brassica napus* and *Arabidopsis thaliana* which can be used by Brassica researchers to identify markers and candidate genes in regions of interest.

Acknowledgements:

AFIF are acknowledged in any communications discussing this project with public and private sector researchers. AFIF was acknowledged in a recent presentation given by Dr Parkin at the 13th Crucifer Genetics Workshop in Davis, CA, March 23-26, 2002, entitled 'Comparative mapping in Crucifers and its role in the genomics age.'

AFIF will be acknowledged in the manuscript in the final stages of preparation describing the global comparative map between *Brassica napus* and *Arabidopsis thaliana*. AFIF will also be acknowledged in a second manuscript describing the development of the *Brassica napus* BAC library and the development of contigs homologous to a region of chromosome five from *A. thaliana*.